

Structure primaire de la caséine κ B bovine

Séquence complète

Jean-Claude MERCIER, Ghislaine BRIGNON et Bruno RIBADEAU-DUMAS

Laboratoire de Recherches sur les Protéines, Institut National de la Recherche Agronomique, Jouy-en-Josas

(Reçu le 5 février 1973)

In a previous report [1], we have given the complete primary structure of κ B₁-caseinomacropptide which is the soluble COOH-terminal fragment split from bovine κ B₁-casein by rennin. We also reported on the COOH-terminal sequence of the NH₂-terminal fragment, the so-called para- κ -casein.

The present paper deals with the complete amino acid sequence of bovine κ B-casein, which has now been achieved by establishing the primary structure of para- κ B-casein of which we discuss the salient features. This work has been reported briefly in a short communication [2]. SCM-para- κ B-casein and maleyl κ BCN1, the NH₂-terminal cyanogen bromide fragment of κ B-casein [1], were used as starting material. The tryptic and peptic peptides of SCM-para- κ B-casein and the chymotryptic peptides of κ BCN1 were isolated on Dowex 50 and Sephadex G-50 or G-25, and their sequences were determined either partially or completely by using classical methods and in some cases mass spectrometry. All these peptides and a NBS fragment of SCM-para- κ B-casein have provided all the overlaps needed for the completion of the amino acid sequence of para- κ B-casein.

Para- κ B-casein is a single polypeptide chain containing 105 amino acid residues: Asp₃, Asn₄, Thr₃, Ser₇, PyroGlu₁, Glu₄, Gln₁₂, Pro₁₂, Gly₁, Ala₉, Val₅, 1/2 Cys₂, Met₁, Ile₆, Leu₇, Tyr₉, Phe₄, Lys₆, His₃, Trp₁, Arg₅, and its molecular weight has been calculated to be 12269. The average hydrophobicity, calculated according to Bigelow [3], is 5.48 kJ (1.310 kcal) per residue, and para- κ B-casein can be therefore considered to be a very hydrophobic molecule. Its net positive charge at pH of native milk (about 6.8) is very close to 4.5. The high content (11.5%) and rather uniform distribution of prolyl residues are incompatible with much α -helical organization of the molecule, as previously shown for κ -casein [4]. Both hydrophobic and charged amino acid residues are distributed non-uniformly along the chain. Two regions, 1–24 and 80–105, are hydrophilic: the very hydrophilic former, with NH₂-terminal pyroglutamic acid, contains a cysteinyl residue located inside a cluster of eleven ionizable residues including six out of the seven total dicarboxylic amino acids; the 80–105 region, which contains the second cysteinyl residue in position 88, is rather hydrophilic, except at the COOH-terminal end which is hydrophobic in spite of the presence of a cluster of four basic residues. These two hydrophilic regions are very likely to be on the outside of the molecule and this may favor the formation of intermolecular S-S bonds. The very hydrophobic central part of the chain, viz., 25–79, where most of the aromatic residues are located, has a para- κ -casein-like behaviour in aqueous solution, and it may be responsible for the aggregation ability of para- κ -casein.

According to the sequence data of both κ B₁-caseinomacropptide [1] and para- κ -casein, it is concluded that bovine κ B-casein is made up of a single polypeptide chain containing 169 amino acid residues: Asp₄, Asn₇, Thr₁₄, Ser₁₂, SerP₁, PyroGlu₁, Glu₁₂, Gln₁₄, Pro₂₀, Gly₂, Ala₁₅, Val₁₁, 1/2 Cys₂, Met₂, Ile₁₃, Leu₃, Tyr₉, Phe₄, Lys₆, His₃, Trp₁, Arg₅, with NH₂-terminal pyroglutamic acid. However, in accordance with the well-known lability of glutamine residues in NH₂-terminal position [5], such a pyroglutamic acid residue may arise from the subsequent cyclisation of a NH₂-terminal glutamine residue present originally in native κ -casein. The content of proline residues is high (11.8%). The molecular weight of the carbohydrate-free monomer of κ B-casein is 19023. Its net negative charge at pH of native milk is close to 3.5 and its average hydrophobicity is about 5.37 kJ (1.285 kcal) per residue.

We have already reported [6] the location of the two amino acid substitutions differentiating the two known genetic variants: the B variant differs from the A variant by the respective amino acid substitutions Ile 136/Thr and Ala 148/Asp.

Tryptic and chymotryptic peptides isolated by Jollès *et al.* [7] from κ A-casein correspond obviously to the fragments 11–16, 17–25, 31–34, 63–67, 1–10 and 86–105 of κ -casein strand, but there is only partial agreement concerning the sequences of amino acid proposed for the last three fragments. More recently, Fiat *et al.* [8] have studied a short glycopeptide isolated from bovine κ -casein, but neither the amino acid sequence nor the location of the phosphate group agree with our results.

Les précédents travaux du laboratoire [1] relatifs à la détermination de la structure primaire de la caséine κ bovine ont permis d'établir la séquence des 75 derniers résidus d'acides aminés de cette phosphoglycoprotéine constituée d'une unique chaîne peptidique de 169 résidus. Ce sont les 64 résidus du caséinomacropéptide, le fragment COOH-terminal résultant de l'hydrolyse par la présure d'une liaison très labile de la caséine κ dans lequel la position exacte des deux substitutions Thr/Ile et Asp/Ala différenciant les variants génétiques A et B a été déterminée [6], et 11 résidus de l'extrémité COOH-terminale du fragment NH₂-terminal appelé paracaséine κ .

Dans le présent travail, qui a fait l'objet d'une brève communication [2], nous avons établi la séquence complète des 105 résidus d'acides aminés de la paracaséine κ (dont Jollès *et al.* [7] avaient donné antérieurement des éléments de structure) en analysant des fragments tryptiques, pepsiques et «NBS» de la SCM-paracaséine κ ainsi que des fragments α -chymotrypsiques du peptide κ BCN1 [1] qui est le fragment «CNBr» NH₂-terminal de la caséine κ B.

Ainsi s'achève l'étude de la structure primaire de la caséine κ bovine dont nous donnons la séquence complète.

MATÉRIEL ET MÉTHODES

Matériel

La caséine κ B a été préparée selon Zittle et Custer [10], mais en omettant le traitement final à l'éthanol, à partir du lait d'une vache de race

Ce travail fait partie d'une Thèse de Doctorat d'Etat (J.-C. Mercier), Université de Paris VII.

Abréviations non usuelles: Met*, méthionine présente sous la forme homosérine-homosérine lactone; HSer, homosérine; SerP, phosphosérine; peptides CN et NBS, peptides obtenus respectivement par action du bromure de cyanogène et de la *N*-bromosuccinimide; SCM- κ , paracaséine κ *S*-carboxyméthylée.

Enzymes: Présure (EC 3.4.4.3); trypsine (EC 3.4.4.4); α -chymotrypsine (EC 3.4.4.5); pepsine (EC 3.4.4.1); thermolysine (EC 3.4.4.-); leucine aminopeptidase (EC 3.4.1.1); aminopeptidase M (EC 3.4.1.-); carboxypeptidase A (EC 3.4.2.1); carboxypeptidase B (EC 3.4.2.2); carboxypeptidase C (EC 3.4.2.-).

française frisonne pie-noire, homozygote pour l'allèle κ -Cn^B. Le fragment «CNBr» NH₂-terminal de la caséine κ B (peptide κ BCN1) a été préparé selon les méthodes précédemment décrites [1]. La paracaséine κ B a été séparée d'un hydrolysate «présure» de la caséine κ B par chromatographie sur carboxyméthylcellulose microgranulaire en tampon Tris-diméthylformamide selon Kim *et al.* [9]; son homogénéité a été vérifiée par électrophorèse sur gel d'amidon selon Schmidt [10].

A l'exception de la présure¹ (Hansen, cristallisée), de la thermolysine (Merck, 6000 PU/mg), de l'aminopeptidase M et de la carboxypeptidase C (Röhm, 2500 mU/mg), les enzymes utilisées proviennent de la Firme Worthington Biochemical Corporation: trypsine (2 $\times$ cristallisée, TRL, 7AA); pepsine (2 $\times$ cristallisée, PM); α -chymotrypsine (3 $\times$ cristallisée, CDI 7CD); leucine aminopeptidase (LAP DFP, 5,9 mg/ml); carboxypeptidase A (COA DFP, 50 mg/ml); carboxypeptidase B (COB DFP, 2,9 mg/ml).

L'origine des autres produits est la suivante: anhydride maléique, Prolabo; isothiocyanate de phényle, Fluka, redistillé; acide trifluoroacétique, Eastman; bromure de cyanogène, Schuchardt; *N*-bromosuccinimide, Baker Organic Chemical; hydrazine hydratée, Prolabo, retraitée au laboratoire [1] (indice de réfraction de l'hydrazine anhydre obtenue: $[\alpha]_D^{25} = 1,467$); carboxyméthylcellulose microgranulaire CM 32, Whatman; Dowex AG-50 W-X2, 200–400 mesh, Biorad; Sephadex, Pharmacia.

Méthodes

Traitement du peptide κ BCN1 par l'anhydride maléique. Le peptide κ BCN1, normalement très insoluble à pH 8,5, se dissout parfaitement bien après blocage des résidus lysyle par l'anhydride maléique [11]. La technique de maléylation que nous avons utilisée est identique à celle appliquée à la caséine α_{s1} [12], mais la réaction a été effectuée en milieu chlorure de guanidine 3 M. Après élimination de l'excès de réactifs par dialyse contre un tampon *N*-éthylmorpholine 0,02 M, pH 8,5, la solution du peptide κ BCN1 maléylé reste parfaitement claire.

¹ Cette préparation nous a été fournie gracieusement par les Etablissements Hansen que nous remercions vivement.

S-Carboxyméthylation de la paracaséine κ B. La SCM-paracaséine κ B, résultant de l'alkylation par l'acide iodoacétique de la paracaséine κ B préalablement réduite en milieu urée et 2-mercaptoéthanol, a été préparée selon Woychik [13].

Hydrolyses enzymatiques. L'hydrolyse de la caséine κ B par la présure a été effectuée dans les conditions de Kim *et al.* [9]. La carboxypeptidase C a été utilisée en tampon citrate de sodium 0,1 M, pH 5,3 selon Sprössler *et al.* [14], mais les doses d'enzymes nécessaires pour dégrader les peptides sont nettement plus élevées que celles indiquées par ces auteurs. Les conditions générales d'emploi des autres enzymes ont été décrites précédemment: trypsine [12]; pepsine, α -chymotrypsine, leucine aminopeptidase et carboxypeptidases A et B [15]; thermolysine [16]; aminopeptidase M [17]. Seules seront donc précisées les conditions particulières: quantité de peptide, rapport enzyme/substrat, température, durée de la réaction, etc. En ce qui concerne les exopeptidases, le volume de la préparation commerciale employé a été rapporté à 0,1 μ mole de peptide, et les résultats ont été exprimés en moles d'acides aminés libérés par mole de peptide.

Hydrolyse par la N-bromosuccinimide. Pour obtenir une coupure sélective au niveau du groupe carbonyle de l'unique résidu tryptophanyle [18] de la SCM-paracaséine κ , nous avons fait agir la N-bromosuccinimide en présence d'acétate de mercure, dans les conditions décrites par Ramachandran et Witkop [19].

Autres techniques. L'hydrazinolyse a été effectuée selon la technique de Fraenkel-Conrat et Chung Ming Tsung [20]. Les autres techniques utilisées dans ce travail ont été décrites en détail ou référencées dans nos précédentes publications: chromatographie sur colonne et sur papier, électrophorèse sur papier, dosage des acides aminés, du tryptophane et du phosphore [12]; hydrolyse par le bromure de cyanogène [21]; dégradation soustractive d'Edman [15].

Spectrométrie de masse. La séquence de quelques peptides a été vérifiée ou précisée par spectrométrie de masse, après N-acétylation et perméthylation ou perdeutériométhylation [22, 23].

Nomenclature des peptides. Les peptides ont été désignés selon le traitement enzymatique ou chimique ayant permis leur obtention; T = trypsine; Th = thermolysine; P = pepsine; C = α -chymotrypsine; CN = bromure de cyanogène; NBS = N-bromosuccinimide; m = désigne un peptide maléylé.

RÉSULTATS

Composition en acides aminés de la paracaséine κ

La composition en acides aminés de la paracaséine κ B (Tableau 1), déduite des résultats de l'hydrolyse acide totale de la SCM-paracaséine κ B

Tableau 1. Composition en acides aminés de la paracaséine κ . Les chiffres gras sont ceux qui diffèrent entre eux

Acide aminé	SCM- κ B ^a	Somme ^{a, b} κ B ₁ CN1 + κ B ₁ CN3 [1]	Séquence
mol d'acides aminés/mol de peptide			
Acide aspartique	6,8	7	7
Thréonine	2,8 ^c	3	3
Sérine	6,4 ^c	7	7
Acide glutamique	16,0	17	17
Proline	11,5	13	12
Glycine	1,1	1	1
Alanine	8,5	9	9
Cystéine	2,2	2	2
Valine	5,0 ^d	5	5
Méthionine	1,2	1	1
Isoleucine	5,5 ^d	6	6
Leucine	7,4	7	7
Tyrosine	8,8 ^c	8	9
Phénylalanine	4,0	4	4
Lysine	6,0	6	6
Histidine	3,1	3	3
Tryptophane	1,1	1 ^e	1
Arginine	5,0	5	5
Phosphore	—	—	—
Nombre de résidus		105	105

^a Sauf indication contraire, les chiffres correspondent à la moyenne des duplicata de 3 temps d'hydrolyse, respectivement 24, 48 et 96 h.

^b La paracaséine κ B correspond à la somme des peptides κ B₁CN1 et κ B₁CN3, si l'on fait abstraction du résidu homosérine-homosérine lactone COOH-terminal de ce dernier [1].

^c Chiffres extrapolés linéairement au temps zéro.

^d Chiffres correspondant à une hydrolyse de 96 h.

^e Valeur donnée par Spies [18].

ou calculée à partir des compositions en acides aminés des deux fragments «CNBr» κ B₁CN1 et κ B₁CN3 de la caséine κ B₁ [1], est dans l'ensemble en bon accord avec celles rapportées antérieurement par divers auteurs [24, 25, 9]. Si l'on se réfère aux acides aminés basiques par exemple, seules diffèrent les estimations du nombre de résidus lysyle qui serait égal à 6 ou 7. En fait, le premier chiffre paraît le plus correct, car il correspond exactement à la différence entre les valeurs de 9 et 3 trouvées pour la lysine, respectivement dans la caséine κ [26, 27, 1] et le caséinomacropéptide [1, 25, 28–32].

PEPTIDES TRYPSIQUES DE LA SCM-PARACASEINE κ B

La paracaséine κ comportant 5 résidus arginyle et 6 résidus lysyle, on pouvait s'attendre à obtenir par hydrolyse tryptique une douzaine de peptides au minimum. En fait, seuls quelques fragments tryptiques ont été obtenus avec un bon rendement, vraisemblablement parce que certaines liaisons Lys/Arg-X sont difficilement attaquées par la trypsine: soit parce qu'elles sont difficilement accessibles à l'enzyme en raison de l'insolubilité de la paracaséine κ dans les conditions de pH de l'hydrolyse, soit parce qu'elles sont du type Lys/Arg-Pro.

La solution de SCM-paracaséine α B dans le tampon *N*-éthyl-morpholine 0,2 M, pH 8,5 étant très hétérogène, nous avons utilisé les conditions d'hydrolyse suivantes: rapport molaire enzyme/substrat = $1/_{100}$; 40 °C; 17 h d'incubation. Le surnageant et le précipité de l'hydrolysats tryptique ont été séparés par centrifugation (20 000 tours/min; 20 min) et analysés séparément.

Etude du surnageant de l'hydrolysats tryptique

Les peptides tryptiques du surnageant ont été séparés par chromatographie sur Dowex 50 (Annexe, Fig.1), puis sur Sephadex G-25. Certains d'entre eux ont été purifiés par électrophorèse ou chromatographie préparative sur papier.

L'annexe, Tableaux 1–6 et la Fig.1 présentent l'ensemble des résultats donnant la séquence partielle ou complète des peptides tryptiques ainsi obtenus.

Peptide T1 (Annexe, Tableau 1) (Asx, Glx₅)-Gln-Pro-Ile-Arg

La séquence partielle COOH-terminale -Gln-Pro-Ile-Arg a été établie à l'aide des carboxypeptidases A, B et C. La séquence NH₂-terminale n'a pu être déterminée car la leucine aminopeptidase et la méthode d'Edman donnent des résultats négatifs. Une analyse plus poussée de la structure a été faite en étudiant des fragments peptiques résultant de l'hydrolyse du peptide T1 par la pepsine (rapport molaire enzyme/substrat = $1/_{50}$; 20 °C; 24 h d'incubation en milieu acide formique 5 vol. %). Trois fragments peptiques ont été séparés à l'état pur par chromatographie sur Dowex 50, et analysés à l'aide des exopeptidases.

La comparaison des éléments de séquence déterminés sur le peptide T1 et ses fragments peptiques, a permis de préciser sa structure, soit:

T1: (Glx, Glx, Glx, Asx, Glx, Glx)-Gln-Pro-Ile-Arg
T1P1: Glx, Glx-(Gln, Asn, Gln)
T1P2: Glx, Glx, Glx, Pro, Ile, Arg
T1P3: (Asn, Gln, Glu)-(Glx, Pro, Ile, Arg)
Séquence: Glx-Glx-Gln-Asn-Gln-Glu-Gln-Pro-Ile-Arg.

Finalement, nous avons fait appel à la spectrométrie de masse pour élucider la nature des deux premiers résidus. Après élimination du résidu arginyle COOH-terminal à l'aide de la carboxypeptidase B, et perdeutériométhylation, le peptide T1 a été analysé au spectromètre de masse. L'étude du spectre obtenu a permis d'établir la nature et l'enchaînement de tous les résidus d'acides aminés constitutifs de l'échantillon injecté, soit [2]:

PyroGlu-Glu-Gln-Asn-Gln-Glu-Gln-Pro-Ile.

Ces résultats ont été confirmés en appliquant la même méthode à l'étude du peptide T1P3.

Peptide T2 (Annexe, Tableau 2) Ser-Cys-Gln-Ala-Gln-Pro-Thr-Thr-Met-Ala-Arg

L'étude du peptide α B, CN3, qui est le fragment «CNBr» central de la caséine α , nous a permis d'établir précédemment [1] la séquence COOH-terminale de la paracaséine α , soit: Met-Ala-Arg-His-Pro-His-Pro-His-Leu-Ser-Phe. Le peptide T2, qui contient l'unique résidu méthionyle de la paracaséine α , possède donc la séquence COOH-terminale -Met-Ala-Arg. Ceci a été vérifié à l'aide des carboxypeptidases A et B qui libèrent, outre ces 3 résidus, un résidu thréonyle. La leucine aminopeptidase libère 4 acides aminés dont une glutamine. La nature du second résidu Glx a été déduite de la mobilité du peptide T2, préalablement hydrolysé par les carboxypeptidases A et B, en électrophorèse sur papier à pH 6,5. Six stades de la dégradation d'Edman ont donné les enchaînements manquants. Grâce au peptide T2, nous connaissons donc l'enchaînement de 19 résidus d'acides aminés de la partie COOH-terminale de la paracaséine α .

Peptide T3 (Annexe, Tableau 3) Ser-Pro-Ala-Gln-(Ile, Leu, Gln, Trp, Glx)-Val- Leu-Ser-Asp-Thr-Val-Pro-Ala-Lys

Les carboxypeptidases A et B libèrent uniquement le résidu lysyle; l'enchaînement des 4 premiers résidus a été déterminé à l'aide de la dégradation soustractive d'Edman. Une analyse plus complète de la séquence a été faite en étudiant des fragments «thermolysine» séparés d'un hydrolysats «thermolysine» du peptide T3 (rapport molaire enzyme/substrat = $1/_{200}$; 40 °C; 24 h d'incubation) par chromatographie sur Dowex 50, et électrophorèse sur papier à pH 1,9.

Fragment NH₂-terminal. T3Th1: Ser-Pro-Ala-Gln. Trois stades de la dégradation d'Edman en ont donné la séquence complète. Le résidu glutaminyle COOH-terminal a été identifié à l'aide de l'aminopeptidase M.

Fragments COOH-terminaux.

T3Th2: Val-Leu-Ser-Asp-Thr-(Val, Pro, Ala)-Lys
T3Th3: Thr-Val-(Pro, Ala)-Lys
T3Th4: Ala-Lys.

Ces 3 peptides, qui possèdent le résidu lysyle COOH-terminal du peptide T3, représentent des fragments de la partie COOH-terminale de celui-ci, ce qui donne un premier recoupement. Cinq stades de dégradation d'Edman donnent la séquence des 5 premiers résidus du peptide T3Th2, le résidu aspartyle ayant été identifié en hydrolysant ce peptide avec l'aminopeptidase M. L'enchaînement des autres acides aminés a été déduit des résultats de 2 stades de la dégradation d'Edman appliquée au peptide T3Th3.

Fragments de la partie centrale de la chaîne. T3Th5: Gln,Leu, et T3Th6: Gln,Trp. Les seuls fragments obtenus ont été ces deux dipeptides, neutres à pH 6,5. L'ensemble de ces résultats nous donne la séquence partielle du peptide T3 indiquée ci-dessus, qui sera élucidée lors de l'étude des peptides chymotrypsiques du fragment «CNBr» κ BCN1.

Peptide T4 (Annexe, Tableau 4)

Ile-Asn-Asn-Gln-Phe-Leu-Pro-Tyr-Pro-Tyr

La leucine aminopeptidase libère 5 acides aminés dont deux asparagines et une glutamine, et la carboxypeptidase A libère une tyrosine. Les résultats de 7 stades de dégradation d'Edman et d'une hydrazinolyse effectuée sur le peptide dont la tyrosine COOH-terminale avait été préalablement libérée par la carboxypeptidase A ont permis d'établir la séquence complète de ce décapeptide.

Peptides T5 et T6 (Annexe, Tableau 4)

Tripeptide T5. Asp-Glu-Arg. L'hydrolyse du peptide T5 par la leucine aminopeptidase libère de l'acide glutamique et de l'acide aspartique, ce dernier occupant la position NH₂-terminale d'après les résultats d'un stade de dégradation d'Edman. La carboxypeptidase B libère de l'arginine. Le tripeptide T5 représente vraisemblablement le fragment COOH-terminal du peptide T6, et ceci a été vérifié par la suite.

Hexapeptide T6. Cys-Glu-Lys-Asp-Glu-Arg. La carboxypeptidase B libère une arginine. La nature des résidus Asx et Glx a été déterminée à l'aide des carboxypeptidases A et B. La séquence de cet hexapeptide a été déduite des résultats de deux temps d'hydrolyse par la leucine aminopeptidase.

Peptide T7 (Annexe, Tableau 5)

Tyr-Ile-Pro-Ile-Gln-Tyr

L'hydrolyse par la carboxypeptidase A libère de la tyrosine et de la glutamine, dont l'enchaînement a été établi par hydrazinolyse. Quatre stades de dégradation d'Edman donnent la structure complète de ce peptide.

Peptide T8 (Annexe, Tableau 5)

Pro-Ala-Ala-Val-Arg

Trois stades de dégradation d'Edman et une hydrolyse par la carboxypeptidase B, ont permis d'établir la séquence.

Peptides T9 et T10 (Annexe, Tableau 5)

T9: Ile-Ala-Lys, et T10: Phe-Phe-Ser-Asp-Lys.

La séquence du tripeptide T9 a été établie à l'aide des exopeptidases.

La structure du peptide T10 a été déduite des résultats de trois stades de dégradation d'Edman, la position COOH-terminale de la lysine ayant été vérifiée à l'aide de la carboxypeptidase B. La présence du résidu aspartyle a été établie en hydrolysant le peptide avec la leucine aminopeptidase.

On peut remarquer que certains peptides, tels T4 et T7, résultent manifestement de coupures non spécifiques vraisemblablement favorisées par l'action prolongée de la trypsine. En réduisant la durée de l'hydrolyse trypsique, on pouvait donc s'attendre à obtenir des fragments de plus grande taille donnant d'utiles recoupements.

La paracaséine κ B a donc été hydrolysée par la trypsine dans les mêmes conditions que précédemment, mais en réduisant la durée d'incubation à 3 h. L'hydrolysate très hétérogène a été desséché, repris en totalité par l'acide formique 75 vol. %, et chromatographié sur Dowex 50. Des diverses fractions obtenues, nous avons isolé par chromatographie sur Sephadex G-25 et électrophorèse sur papier, la plupart des peptides précédemment étudiés, ainsi que 3 nouveaux peptides, dénommés T11, T12 et T13.

Peptides T11 et T12 (Annexe, Tableau 6)

T11: Val-Leu-Ser-Arg
T12: Tyr-Ile-Pro-Ile-Gln-Tyr-Val-Leu-Ser-Arg

Les résultats de l'hydrolyse par les carboxypeptidases A et B, et de deux stades de dégradation d'Edman, donnent la séquence complète du tétrapeptide T11. La composition en acides aminés du peptide T12 est la somme de celles des peptides T11 et T7; l'enchaînement T7-T11 a été établi en analysant la séquence COOH-terminale du peptide T12 à l'aide des carboxypeptidases A et B, et en vérifiant la présence d'un résidu NH₂-terminal tyrosyle à l'aide de la leucine aminopeptidase.

Peptide T13 (Annexe, Tableau 6)

Phe-Phe-Ser-Asp-Lys-Ile-Ala-Lys

La composition en acides aminés de cet octapeptide est la somme de celles des peptides T9 et T10 dont l'enchaînement T10-T9 a été déduit des résultats de l'hydrolyse du peptide T13 par la leucine aminopeptidase et les carboxypeptidases A et B.

A ce stade de notre étude, nous connaissons donc la séquence de 13 fragments de la paracaséine κ , représentant environ 80% de la chaîne peptidique (Fig. 1). Le peptide T1 est probablement le fragment NH₂-terminal mais la présence du résidu pyroglutamyle n'est pas une preuve suffisante car il peut résulter de la cyclisation d'un résidu glutaminyle se retrouvant en position NH₂-terminale à la suite de la rupture d'une liaison X-Gln. Enfin, comme nous l'avons fait remarquer précédemment, le peptide T2

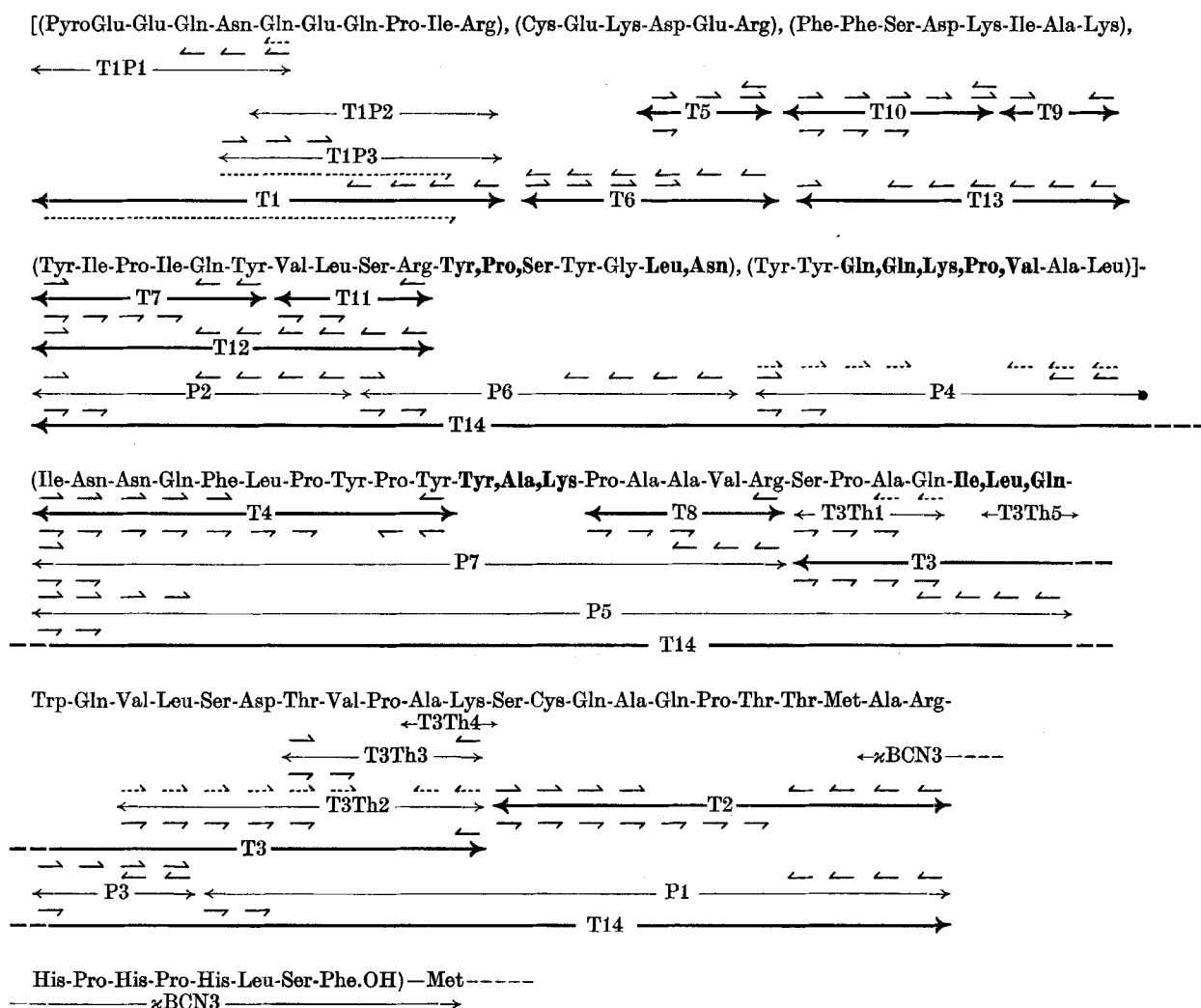


Fig. 1. Séquence partielle de la paracaséine α , déduite des résultats de l'analyse de peptides trypsiques et pepsiques de la paracaséine α . Les acides aminés en caractères gras sont ceux dont l'ordre n'est pas connu. Symboles: acides aminés libérés: → par la leucine aminopeptidase; zzz par l'aminopeptidase M; ← par les carboxypeptidases A, B et C. Séquence établie: → par la méthode d'Edman; ← par hydrazinolysé; --- par spectrométrie de masse. ← T → peptides trypsiques; ← P → peptides pepsiques

est situé en pénultième position par rapport au fragment COOH-terminal de la paracaséine α dont la séquence est His-Pro-His-Pro-His-Leu-Ser-Phe·OH [1]. Ce fragment, où sont localisées les 3 histidines de la paracaséine α , a effectivement été détecté dans le surnageant à l'aide du réactif de Pauly, mais il est resté fixé sur la résine Dowex 50 lors de la chromatographie.

Etude du précipité de l'hydrolysate trypsique

Le précipité de l'hydrolysate trypsique contenait principalement un très gros fragment que nous avons essayé de purifier par chromatographie sur Sephadex G-50 en milieu acétique 60 vol.-%, et dont

nous avons déterminé la composition en acides aminés (Tableau 2). Ce peptide, que nous appellerons T14, étant très insoluble à pH 8,5 mais soluble dans l'acide formique 5 vol.-%, nous avons choisi de l'hydrolyser par la pepsine dans des conditions très ménagées: rapport molaire enzyme/substrat 1/6000; 20 °C; 8 h d'incubation dans l'acide formique 5 vol.-%. L'hydrolysate a été chromatographié sur Dowex 50 (Annexe, Fig. 2) et les fractions obtenues ont été purifiées sur Sephadex G-50 ou G-25; dans quelques cas, une repurification sur papier, par électrophorèse ou chromatographie préparative, a été nécessaire.

Sept peptides pepsiques dont les compositions en acides aminés sont données dans le Tableau 2

Tableau 2. Composition en acides aminés des fragments peptiques du peptide tryptique T14

Entre parenthèses: nombres entiers proposés. Les valeurs de la sérine et de la thréonine, acides aminés partiellement détruits durant l'hydrolyse, n'ont pas été corrigées

Acide aminé	P1	P2	P3	P4	P5	P6	P7	Somme P1 → P6*	T14
mol d'acides aminés/mol de peptide									
Acide aspartique	1,14 (1)	—	—	—	1,96 (2)	1,02 (1)	1,83 (2)	4	4,30 (4)
Thréonine	2,64 (3)	—	—	—	—	—	—	3	2,70 (3)
Sérine	1,59 (2)	—	—	—	0,95 (1)	1,63 (2)	—	5	4,21 (5)
Acide glutamique	2,42 (2)	1,02 (1)	1,19 (1)	2,05 (2)	3,01 (3)	—	1,10 (1)	9	9,01 (9)
Proline	1,94 (2)	1,00 (1)	—	1,02 (1)	4,17 (4)	0,94 (1)	2,85 (3)	9	8,91 (9)
Glycine	—	—	—	—	—	1,05 (1)	—	1	1,17 (1)
Alanine	2,78 (3)	—	—	1,08 (1)	3,92 (4)	—	2,86 (3)	8	7,98 (8)
Cystéine	0,53 (1)	—	—	—	—	—	—	1	0,95 (1)
Valine	0,92 (1)	0,98 (1)	0,89 (1)	0,96 (1)	1,04 (1)	—	0,96 (1)	5	4,63 (5)
Méthionine	0,53 (1)	—	—	—	—	—	—	1	0,64 (1)
Isoleucine	—	1,93 (2)	—	—	1,84 (2)	—	1,07 (1)	4	3,31 (4)
Leucine	—	1,04 (1)	0,91 (1)	0,94 (1)	2,01 (2)	1,03 (1)	1,20 (1)	6	5,48 (6)
Tyrosine	—	1,64 (2)	—	1,78 (2)	2,70 (3)	1,79 (2)	2,63 (3)	9	7,62 (8)
Phénylalanine	—	—	—	—	1,00 (1)	—	1,00 (1)	1	1,39 (1)
Lysine	0,93 (1)	—	—	0,97 (1)	1,04 (1)	—	0,90 (1)	3	3,19 (3)
Histidine	—	—	—	—	—	—	—	—	—
Tryptophane	—	—	0,75 (1)	—	—	—	—	1	—(1)
Arginine	1,00 (1)	—	—	—	0,92 (1)	0,95 (1)	0,93 (1)	3	2,73 (3)
Nombre de résidus	18	8	4	9	25	9	18	73	72

* Le peptide peptique P7 est un fragment du peptide P5. Ces chiffres correspondent à une analyse de 24 h.

ont été obtenus à l'état pur. En faisant abstraction du peptide P7 qui semble être un fragment du peptide P5, hypothèse que nous avons vérifiée ultérieurement, la composition en acides aminés du peptide T14 obtenue par reconstitution à partir de ses fragments peptiques est identique à celle déterminée par analyse directe. L'analyse des extrémités NH_2 - et COOH -terminales de ces 7 fragments peptiques, dont les résultats sont consignés dans l'Annexe, Tableau 7, nous a permis de préciser la localisation de certains peptides tryptiques étudiés précédemment (Fig. 1).

Peptide P1

Ser-Asx-(Thr,Val,Pro,Ala,Lys,Ser,Cys,Glx,
Ala,Glx,Pro,Thr)-Thr-Met-Ala-Arg

Les carboxypeptidases A et B libèrent dans l'ordre Arg,Ala,Met et Thr, et deux stades de dégradation d'Edman donnent Ser-Asx- comme séquence NH_2 -terminale. On remarquera que le peptide P1 a la même séquence COOH -terminale que le fragment tryptique T2, et qu'il contient les 3 résidus thréonyle de la paracaséine α localisés dans les fragments tryptiques T2 et T3. On en déduit que le peptide P1 est constitué dans l'ordre d'un fragment COOH -terminal du peptide T3 et du peptide T2. Ce recoupe-ment nous donne l'enchaînement T3—T2.

Peptide P3

Trp-Gln-Val-Leu

La leucine aminopeptidase libère entre autres un résidu glutaminyle.

La séquence de ce térapeptide a été déduite des résultats d'un stade de dégradation d'Edman, et de deux temps d'hydrolyse par la carboxypeptidase A. Il permet de donner la position exacte du résidu tryptophanyle dans la chaîne peptidique du peptide T3.

Peptide P7

Ile-Asn-(Asn,Gln,Phe,Leu,
Pro,Tyr,Pro,Tyr,Tyr,Ala,Lys,Pro,Ala)-
(Ala,Val)-Arg

La séquence COOH -terminale partielle (Ala,Val)-Arg, déterminée à l'aide des carboxypeptidases A et B, est celle du peptide tryptique T8. La leucine aminopeptidase libère entre autres deux résidus asparaginyne et un résidu glutaminyle. Deux stades de dégradation d'Edman donnent l'enchaînement NH_2 -terminal Ile-Asn qui est celui du peptide T4.

Le peptide P7 serait donc constitué, dans l'ordre, du peptide T4, d'un fragment de composition Tyr, Ala,Lys, et du peptide T8.

Peptide P5

Ile-Asn-(Asn,Gln,Phe,Leu,Pro,Tyr,Pro,Tyr,
Tyr,Ala,Lys,Pro,Ala,Ala,Val,Arg,Ser,Pro,Ala)-
(Gln,Ile,Leu,Gln)

D'après les résultats de l'hydrolyse par la leucine aminopeptidase et de deux stades de dégradation d'Edman, ce peptide a la même séquence NH_2 -terminale que le peptide P7. Le peptide P5 correspond donc au peptide P7 prolongé du côté COOH -terminal

par un fragment dont la composition en acides aminés se déduit par différence de celles des peptides P5 et P7. L'hydrolyse par la carboxypeptidase A libère deux résidus glutaminyle, un résidu leucyle et un résidu isoleucyle. On pourra remarquer que ce fragment COOH-terminal correspond exactement au fragment NH₂-terminal du peptide T3, d'où les recoupements: T4 — (Tyr,Ala,Lys) — T8—T3.

Peptide P4

Tyr-Tyr-(Gln,Gln,Lys,Pro,Val)-Ala-Leu

La carboxypeptidase A libère dans l'ordre la leucine et l'alanine. La séquence NH₂-terminale a été déduite d'un temps d'hydrolyse par la leucine aminopeptidase, et de deux stades de dégradation d'Edman. Ce peptide ne donne aucun recoupement compte tenu des données actuelles.

Peptide P6

Ser-Arg-(Tyr,Pro,Ser)-Tyr-Gly-(Leu,Asn)

La séquence COOH-terminale partielle a été établie à l'aide de la carboxypeptidase A; la séquence NH₂-terminale Ser-Arg- déterminée à l'aide de la méthode de dégradation d'Edman est identique à la séquence COOH-terminale du peptide T12, ce qui nous donne l'enchaînement: T12-(Tyr,Pro-Ser)-Tyr-Gly-(Leu,Asn).

Peptide P2

Tyr-Ile-(Pro,Ile)-(Gln,Tyr,Val,Leu)

Deux stades de dégradation d'Edman donnent l'ordre des deux premiers résidus; la carboxypeptidase A libère 4 résidus. Ce fragment correspond manifestement au peptide tryptique T12 amputé du fragment COOH-terminal Ser-Arg.

Ces résultats, joints à ceux obtenus précédemment lors de l'étude des peptides tryptiques, permettent d'établir la séquence partielle de la paracaséine α (Fig.1). Cette figure donne, à quelques exceptions près, l'enchaînement des résidus d'acides aminés de la moitié COOH-terminale de la chaîne peptidique, ainsi que la séquence complète ou partielle de 5 fragments de la moitié NH₂-terminale dont l'enchaînement reste à établir.

L'étape suivante de notre travail a donc consisté à déterminer les positions respectives des cinq fragments T1, T6, T13, P2—P6 et P4 dans la moitié NH₂-terminale de la chaîne peptidique, et à établir l'enchaînement de quelques résidus d'acides aminés non encore ordonnés. Dans ce but, nous avons étudié les fragments chymotrypsiques provenant du fragment «CNBr» NH₂-terminal de la caséine α B (peptide α BCN1 [1]).

FRAGMENTS CHYMOTRYPSIQUES DU PEPTIDE α BCN1

Les peptides α -chymotrypsiques résultant de l'hydrolyse du peptide α BCN1 (rapport molaire enzyme/substrat 1/200; pH 8,5; 40 °C; 5 h d'incubation), préalablement maléylé pour augmenter sa solubilité [11], ont été séparés par chromatographie sur Dowex 50 (Annexe, Fig.3) et sur Sephadex G-25. Dans quelques cas, une repurification sur papier par électrophorèse à pH 1,9 s'est révélée nécessaire. Nous avons ainsi isolé un grand nombre de peptides chymotrypsiques dont l'analyse a permis d'achever l'étude de la structure primaire de la paracaséine α . Le Tableau 8, en Annexe, présente l'ensemble de nos résultats dans un ordre permettant de suivre l'établissement progressif de la séquence à partir de l'extrémité NH₂-terminale. Ils sont schématisés sur la Fig.2.

Peptide C10

(Glx,Glx,Glx,Asx,Glx,Glx,Glx,Pro,Ile)-

(Arg,Cys,Glu,Lys,Asp,Glu)-Arg-Phe-Phe

La séquence COOH-terminale -Arg-Phe-Phe a été établie en faisant agir à pH 8,5 successivement la carboxypeptidase A puis la carboxypeptidase B. Le même mélange d'enzymes agissant à pH 5,4 libère six résidus supplémentaires. Ni la méthode d'Edman, ni la leucine aminopeptidase n'ont permis d'identifier le résidu NH₂-terminal. En se reportant à la Fig.1, il est évident que le peptide C10 est constitué, dans l'ordre, des peptides T1, T6 et du fragment NH₂-terminal Phe-Phe du peptide T13.

Peptide C17

Ser-Asx-Lys-Ile-Ala-Lys-Tyr

La carboxypeptidase A libère un résidu tyrosyle. Cinq stades de dégradation d'Edman donnent la séquence complète de ce peptide qui correspond au peptide T13 dépourvu de son extrémité NH₂-terminale Phe-Phe, mais possédant un résidu supplémentaire tyrosyle à son extrémité COOH-terminale. Ceci prouve que le fragment consécutif au peptide T13 comporte un résidu NH₂-terminal tyrosyle, ce qui est le cas des deux fragments P2—P6 et P4 dont les positions respectives restent indéterminées à ce stade de notre étude.

Peptide C6

Ile-Pro-Ile-Gln-Tyr

La carboxypeptidase A libère dans l'ordre une tyrosine et une glutamine, la position COOH-terminale du résidu tyrosyle ayant été vérifiée par hydrazinolyse. La séquence complète a été déduite des résultats de trois stades de dégradation d'Edman. Ce peptide C6 correspond au peptide tryptique T7 dépourvu du résidu tyrosyle NH₂-terminal.

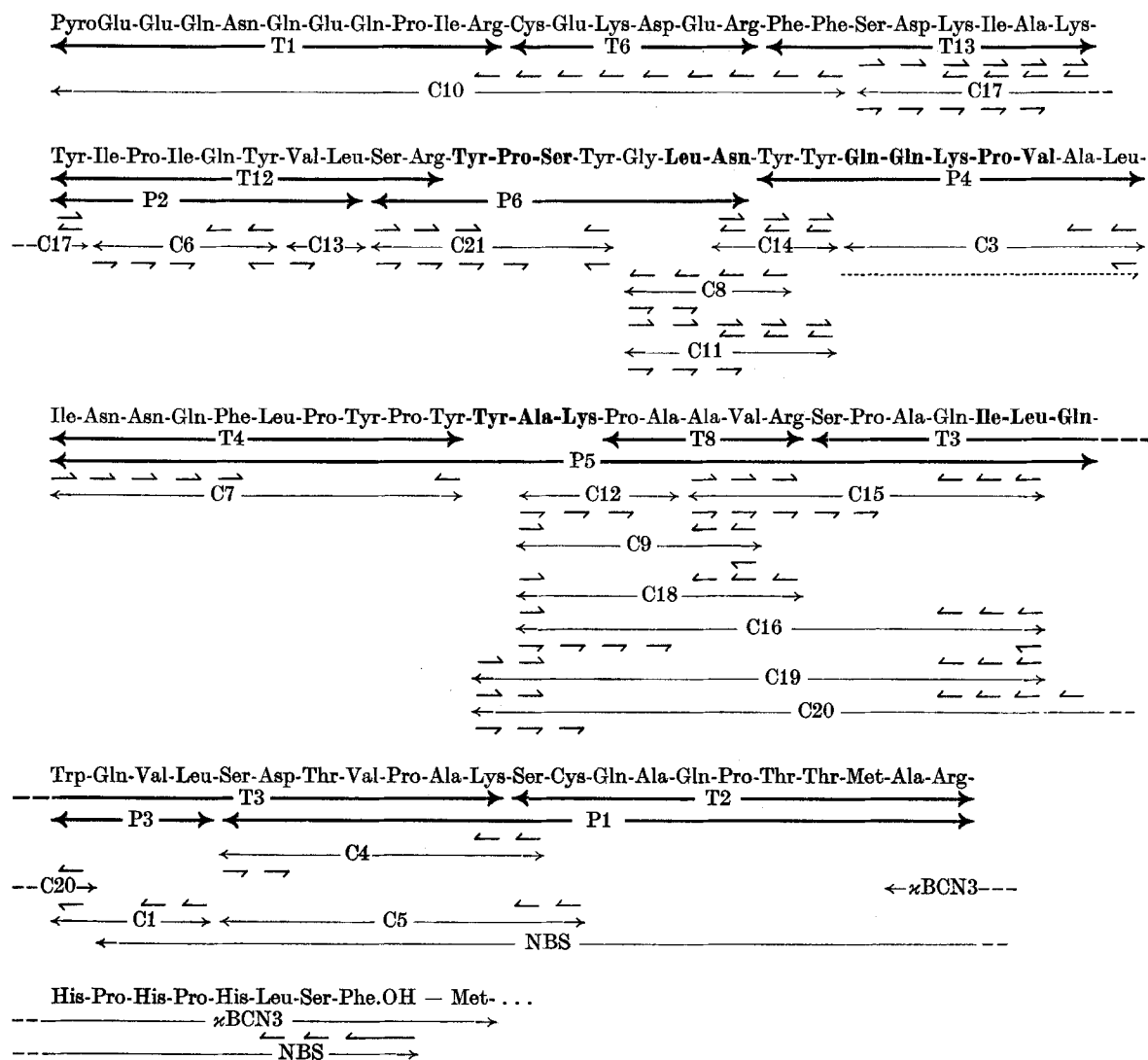


Fig. 2. Structure primaire de la paracaséine α déduite de la Fig. 1 après analyse des peptides α -chymotrypsiques du peptide α BCN1 maléylé. Les acides aminés en caractères gras sont ceux dont l'ordre n'était pas connu avant l'étude des peptides chymotrypsiques. Seuls les principaux peptides trypsiques et pepsiques ont été indiqués; se reporter à la Fig. 1. Symboles: Se reporter à la légende de la Fig. 1 ← T—P → Peptides trypsiques ou pepsiques; ← C → Peptides chymotrypsiques

Peptide C13

Val-Leu

Un stade de dégradation d'Edman donne la séquence de ce dipeptide qui correspond au fragment NH_2 -terminal du peptide trypsique T11.

Peptide C21

Ser-Arg-Tyr-Pro-Ser-Tyr

Quatre stades de dégradation d'Edman et une hydrolyse par la carboxypeptidase A donnent la séquence complète. Ce peptide représente le fragment NH_2 -terminal du peptide pepsique P6 dont la séquence est ainsi pratiquement achevée.

Peptides C8, C11 et C14

C8: Gly-Leu-Asn-Tyr

C11: Gly-Leu-Asn-Tyr-Tyr

C14: Asn-Tyr-Tyr

L'hydrolyse du peptide C11 par la carboxypeptidase A libère dans l'ordre deux résidus tyrosyle et un résidu asparaginylo. Deux stades de dégradation d'Edman donnent l'enchaînement complet des acides aminés de ce peptide, confirmé par l'analyse des peptides C8 et C14. Les deux résidus tyrosyle sont manifestement communs aux peptides C11 et P4, ce qui nous donne le recoupement (P2—P6) — P4. On remarquera en effet que l'unique

résidu glycyle de la paracaséine α est localisé dans le peptide C11 et le peptide P6.

Nous connaissons donc maintenant l'enchaînement de tous les fragments localisés dans la moitié NH_2 -terminale de la chaîne peptidique de la paracaséine α , soit: T1—T6—T13—P2—P6—P4.

Peptide C3

Gln²-Gln-Lys-Pro-Val-Ala-Leu

Ce peptide correspond au peptide pepsique P4 dépourvu, à son extrémité NH_2 -terminale, des deux résidus tyrosyle. La carboxypeptidase A libère effectivement de l'alanine et de la leucine dont l'enchaînement Ala-Leu a été établi par hydrazinolyse. Les résultats négatifs obtenus avec la leucine aminopeptidase et la méthode d'Edman suggèrent la présence d'un résidu NH_2 -terminal pyroglutamyle résultant de la cyclisation du résidu glutaminyle en position NH_2 -terminale après hydrolyse de la liaison Tyr-Gln. Ceci a été confirmé en analysant au spectromètre de masse la séquence complète du peptide C3 préalablement perméthylé [2].

Peptide C7

Ile-Asn-Asn-Gln-Phe-Leu-Pro-Tyr-Pro-Tyr

Ce peptide est identique au peptide tryptique T4 étudié précédemment (Annexe, Tableau 4).

Peptides C19 et C16

C19: Tyr-Ala-(Lys,Pro,Ala,Ala,Val,Arg,Ser,Pro,Ala)-(Gln,Ile,Leu)

C16: Ala-Lys-Pro-Ala-(Ala,Val,Arg,Ser,Pro,Ala)-Gln-Ile-Leu

L'hydrolyse de ces deux peptides par la carboxypeptidase A libère les trois mêmes résidus leucyle, isoleucyle et glutaminyle dont l'enchaînement a été déduit des résultats de plusieurs temps d'hydrolyse. Le résidu supplémentaire tyrosyle du peptide C19 se trouve donc en position NH_2 -terminale, ce qui a été vérifié à l'aide de la leucine aminopeptidase. Quatre stades de dégradation d'Edman effectués sur le peptide C16 donnent la séquence des quatre premiers résidus.

Ces deux peptides permettent de vérifier l'enchaînement (Tyr-Ala-Lys) — T7 — T13 déduit de l'étude des peptides pepsiques P7 et P5, et donnent l'enchaînement des derniers résidus non encore ordonnés de la paracaséine α .

Peptide C20

Tyr-Ala-Lys-(Pro,Ala,Ala,Val,Arg,Ser,Pro,Ala)-(Gln,Ile,Leu)-Gln-Trp

Ce peptide diffère du peptide C19 par la présence des résidus glutaminyle et tryptophanyle. La position

² sous forme pyroGlu.

COOH -terminale de ce dernier a été établie par hydrazinolyse. Outre ce résidu, la carboxypeptidase A libère la leucine, l'isoleucine et deux glutamines. La leucine aminopeptidase libère la tyrosine et l'alanine. Trois stades de dégradation d'Edman confirment l'enchaînement Tyr-Ala-Lys.

Peptides C12, C9 et C18

C12: Ala-Lys-Pro-Ala

C9: Ala-(Lys,Pro,Ala)-Ala-Val

C18: Ala-(Lys,Pro,Ala)-(Ala,Val)-Arg

L'hydrolyse de ces trois peptides par la leucine aminopeptidase libère un résidu alanyl. Trois stades de dégradation d'Edman donnent la séquence complète du peptide C12. Le peptide C18 diffère du peptide C9 par la présence du résidu supplémentaire arginyle dont la position COOH -terminale a été vérifiée à l'aide de la carboxypeptidase B.

Ces trois peptides confirment l'enchaînement (Ala-Lys) — T7.

Peptide C15

Ala-Val-Arg-Ser-Pro-Ala-Gln(Ile,Leu)

L'enchaînement des cinq premiers résidus a été établi à l'aide de la dégradation d'Edman. La carboxypeptidase A libère la leucine, l'isoleucine et la glutamine. Ce peptide permet de vérifier l'enchaînement T7—T3 établi précédemment.

Peptides C1 et C2

C2: Gln,Trp

C1: (Trp,Gln)-Val-Leu

Le tétrapeptide C1 est identique au peptide pepsique P3.

Peptides C4 et C5

C4: Ser-Asx-(Thr,Val,Pro,Ala)-Lys-Ser

C5: Ser-Asx-(Thr,Val,Pro,Ala)-Lys-Ser-Cys

Deux stades de dégradation d'Edman donnent la séquence des deux premiers résidus. Les deux peptides ayant la même séquence NH_2 -terminale, on en déduit que le résidu cystéyle est situé en position COOH -terminale dans le peptide C5, ce qui a été vérifié à l'aide de la carboxypeptidase A qui libère un résidu cystéyle et un résidu sérile. Ces deux peptides confirment l'enchaînement T3—T2.

En résumé, l'étude des fragments chymotrypsiques du peptide αBCN1 nous a permis d'établir les enchaînements manquants (se reporter à la Fig. 1) et de vérifier en grande partie ceux qui avaient été déduits de l'analyse des peptides tryptiques et pepsiques.

PEPTIDE «NBS» DE LA SOM-PARACASEINE αB

Nous avons également analysé un fragment «NBS» de la paracaséine α , isolé par chromato-

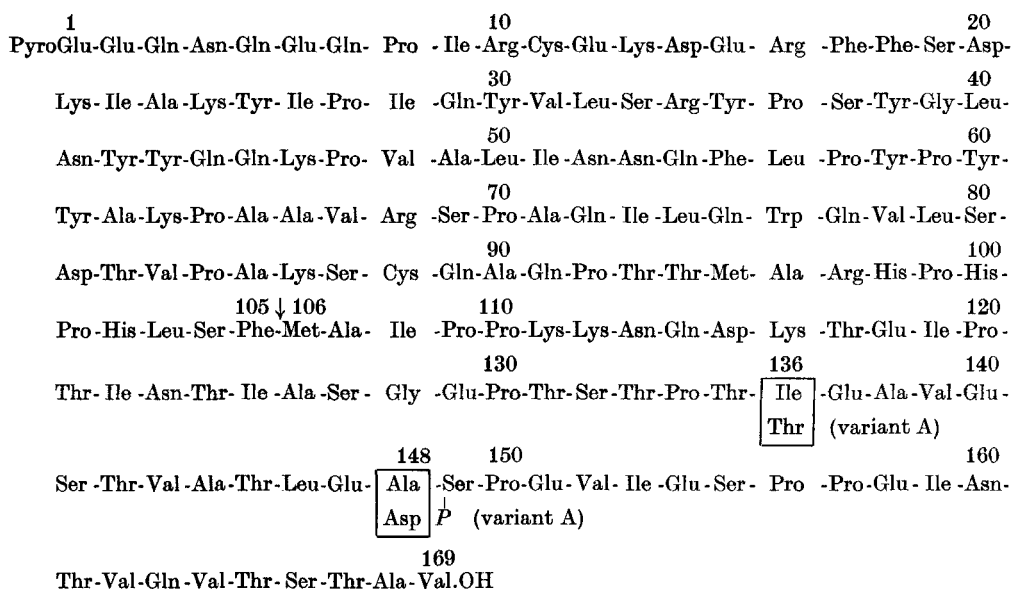


Fig. 3. Structure primaire de la caséine α B bovine. Les résidus d'acides aminés encadrés sont les sites correspondant à la double mutation différenciant le variant génétique A du variant de référence B [6]. $\downarrow$ Liaison peptidique hydrolysée par la présure

graphies successives sur Sephadex G-50, à partir d'un hydrolysate de la SCM-paracaséine α B par la N-bromosuccinimide. Dans les conditions utilisées [19], on pouvait s'attendre à obtenir deux fragments résultant de l'hydrolyse sélective de la liaison Thr-Gln:

a) Un fragment COOH-terminal ayant pour composition en acides aminés Asx, Thr₃, Ser₃, Glx₃, Pro₄, Ala₃, Val₂, $\frac{1}{2}$ Cys, Met, Leu, Phe, Lys, His₃, Arg, et comme séquences NH₂- et COOH-terminales, respectivement Gln-Val-Leu- et -Leu-Ser-Phe-OH. Ce sont précisément les caractéristiques du peptide que nous avons isolé (Annexe, Tableau 9); notons cependant que nous n'avons pas été en mesure d'identifier le résidu NH₂-terminal à l'aide de la leucine aminopeptidase, mais cet insuccès suggère la présence d'un résidu glutaminyle, sous forme cyclisée, à l'extrémité NH₂-terminale.

b) Un fragment NH₂-terminal que nous n'avons pu séparer de la SCM-paracaséine α B résiduelle dont la présence était due à une hydrolyse incomplète.

Quoi qu'il en soit, le fragment «NBS» que nous avons obtenu confirme la séquence COOH-terminale à partir du résidu tryptophanyle.

DISCUSSION

La paracaséine α est constituée d'une unique chaîne peptidique de 105 résidus d'acides aminés, bloquée à son extrémité NH₂-terminale par un résidu pyroglutamyle (Fig. 3). Sa composition en acides aminés Asp₃, Asn₄, Thr₃, Ser₇, PyroGlu₁, Glu₄, Gln₁₂, Pro₁₂, Gly₁, Ala₉, Val₅, $\frac{1}{2}$ Cys₂, Met₁,

Ile₆, Leu₇, Tyr₉, Phe₄, Lys₆, His₃, Trp₁, Arg₅, déduite de la séquence, concorde avec celle que nous avons établie précédemment en dosant les acides aminés respectivement dans la SCM-paracaséine α B et dans les fragments «CNBr» de la caséine α B₁ [1] (Tableau 1). Elle est très proche de celles rapportées antérieurement par plusieurs auteurs [9, 24, 25], les divergences portant généralement sur les acides aminés dont la teneur est élevée.

Les travaux du groupe de Jollès mis à part, ceux qui ont été consacrés à la structure primaire de la paracaséine α se sont limités en fait à l'étude des extrémités NH₂- et COOH-terminales. L'échec de toutes les tentatives faites jusqu'à présent pour déterminer la séquence NH₂-terminale, laissait supposer l'existence d'un résidu particulier et la présence d'un résidu pyroglutamyle avait été suggérée mais non démontrée. La séquence COOH-terminale -Leu-PheOH suggérée initialement [33, 34] avait été ensuite infirmée et le nouvel enchaînement -Leu-Ser-PheOH proposé [35] devait être vérifié par la suite [36, 37]. Ultérieurement, Jollès *et al.* établissaient la séquence de 7 [36] puis de 18 résidus [7] de l'extrémité COOH-terminale. La séquence du fragment 87-105 (Fig. 3), correspondant à cette partie de la chaîne, est pratiquement identique à celle proposée par Jollès *et al.* [7] dont elle ne diffère que par la présence d'un troisième résidu histidyle. Cependant, comme nous l'avons fait remarquer précédemment [1], cette différence apparemment minime peut avoir des répercussions importantes dans le domaine de la recherche de substrats synthétiques de la présure. Les autres éléments de

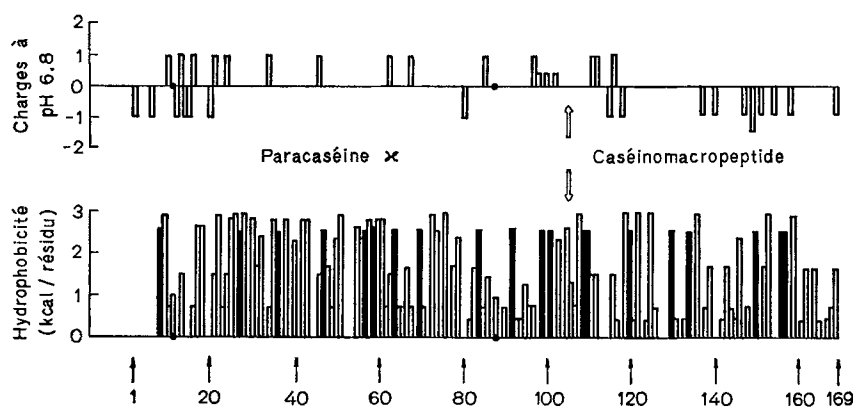


Fig.4. Diagramme montrant la répartition des résidus hydrophobes et des charges électriques dans la chaîne peptidique de la caséine α B. ↓ Liaison hydrolysée par la présure; ● résidu cystéyle. ■ résidu prolyle. Les valeurs de l'hydrophobicité utilisées pour les diverses chaînes latérales des résidus d'acides aminés sont celles de Bigelow [3]

séquence donnée par Jollès *et al.* [7] correspondent manifestement aux fragments 11–16, 17–25, 31–34, 63–67, 1–10 et 25–30 de la chaîne peptidique de la paracaséine α . Toutefois, les séquences proposées pour les deux derniers fragments (respectivement Gln-Gln-Gln-Asn-Glu-Glu-Glu-Pro-Ile-Arg et Tyr-Ile-Pro-Ile-Tyr-Gln) diffèrent de celles que nous avons établies.

La masse moléculaire du monomère, calculée d'après la séquence, est égale à 12269 daltons. Son hydrophobicité moyenne, calculée selon Bigelow [3] est de l'ordre de 5,48 kJ (1,31 kcal) par résidu. Cette valeur élevée contraste avec celle beaucoup plus faible du caséinomacropeptide, qui est de l'ordre de 4,53 kJ (1,083 kcal) par résidu [1].

La teneur élevée en résidus prolyle (11,5%) et leur distribution relativement uniforme au sein de la chaîne peptidique limitent les possibilités de structure en hélice α , ce qui confirme les résultats de Herskovits [4].

La paracaséine α est caractérisée par la prédominance des résidus basiques, et au pH normal du lait, qui est de l'ordre de 6,8, elle possède une charge nette positive voisine de 4,5, ce qui explique sa migration vers la cathode à pH alcalin. Du fait de l'inégale répartition des résidus chargés et des résidus hydrophobes, certaines parties de la chaîne peptidique sont très caractéristiques, et ceci apparaît clairement sur la Fig.4. La région COOH-terminale délimitée par les résidus 80 et 105 est une zone plutôt hydrophile, sauf à son extrémité COOH-terminale à caractère hydrophobe malgré la présence d'un groupe de 4 résidus basiques. La deuxième zone hydrophile de la molécule, correspondant à la partie NH_2 -terminale de la chaîne, est certainement la partie la plus remarquable de la paracaséine α : plus de la moitié des résidus chargés y sont localisés, et, sur les 24 premiers résidus, on dénombre 11 résidus ionisables, dont 6 résidus dicarboxyliques sur les

7 que comporte la molécule; enfin, son extrémité est bloquée du fait de la présence du résidu NH_2 -terminal pyroglutamyle. Quant à la région centrale 25–79 de la chaîne peptidique, c'est une zone très hydrophobe contenant la presque totalité des résidus aromatiques ainsi que quelques résidus basiques uniformément répartis. Cette partie de la chaîne peptidique joue certainement un rôle majeur dans la polymérisation des molécules de paracaséine α en solution aqueuse, si l'on se réfère aux propriétés du peptide T14. Ce peptide tryptique, qui correspond au fragment 25–97, est en effet très insoluble à pH 8,5 et comme la paracaséine α il précipite en formant des fibrilles.

La présence de 2 résidus cystéyle situés respectivement à proximité des extrémités NH_2 et COOH-terminales, est une autre caractéristique de la molécule. Contrairement au résidu cystéyle 11, qui est localisé dans un groupe de 6 résidus chargés, le résidu cystéyle en position 88 est situé dans une zone hydrophile peu ionisée. Ainsi, même si ces deux zones viennent en contact, ce qui ne peut être exclu en raison de la structure relativement lâche de la molécule de caséine α en solution [4,38,39], les interactions électrostatiques et hydrophobes seront suffisamment faibles pour ne pas favoriser la formation d'une liaison disulfure intramoléculaire qui n'a d'ailleurs jamais été mise en évidence. Les deux régions hydrophiles comportant les 2 résidus cystéyle sont probablement situées à l'extérieur de la molécule, ce qui est compatible avec l'existence de liaisons disulfure intermoléculaires.

CONCLUSION

La caséine α B bovine, dont nous avons étudié séparément les deux éléments constitutifs, le caséinomacropeptide [1] et la paracaséine α , est formée d'une seule chaîne peptidique de 169 résidus d'acides

aminés, bloquée à son extrémité NH_2 -terminale par un résidu pyroglutamyle. Cependant, bien que ce résidu ait été identifié formellement par spectrométrie de masse [2], cela ne prouve pas nécessairement la présence de ce résidu dans la caséine κ native. La grande labilité des résidus glutaminyle NH_2 -terminaux est un fait bien connu [5]: leur conversion en résidu pyroglutamyle, qui résulte d'une déamidation suivie d'un réarrangement moléculaire donnant la forme lactame, peut se produire dans des conditions de milieu relativement douces. On ne peut donc exclure la présence d'un résidu NH_2 -terminal glutaminyle dans la caséine κ native, car les conditions d'isolement et de purification de cette protéine sont suffisamment drastiques pour favoriser la réaction de cyclisation.

La composition en acides aminés de la caséine κ B: Asp₄, Asn₇, Thr₁₄, Ser₁₂, SerP₁, Glu₁₂, Gln₁₄, PyroGlu₁, Pro₂₀, Gly₂, Ala₁₅, Val₁₁, $\frac{1}{2}$ Cys₂, Met₂, Ile₁₃, Leu₈, Tyr₉, Phe₄, Lys₉, His₃, Trp₁, Arg₅, déduite de la séquence, est très proche de celle que nous avons calculée en faisant la somme des différents résidus d'acides aminés constitutifs des 3 peptides «CNBr» de la caséine κ B₁ [1] et de celles rapportées antérieurement par divers auteurs [26, 27].

La masse moléculaire du monomère ne contenant pas de sucres, calculée d'après la séquence, est égale à 19023 daltons. Son hydrophobicité moyenne, calculée d'après Bigelow [3], est de l'ordre de 5,37 kJ (1,28 kcal) par résidu. Cette valeur est à comparer aux valeurs de 4,89 kJ (1,17 kcal) et 5,56 kJ (1,33 kcal) par résidu trouvées respectivement pour les caséines α_{s1} B et β A² bovines [40, 17]. L'absence presque complète de structure secondaire, déduite par Herskovits [4] de mesures de dispersion rotatoire optique, est corroborée par la forte teneur (11,8%) et la distribution relativement uniforme des résidus prolyle dans la chaîne peptidique. Au pH normal du lait, qui est de l'ordre de 6,8, le monomère ne contenant pas de sucre de la caséine κ B a une charge nette négative voisine de 3,5.

L'inégale répartition des résidus chargés et des résidus hydrophobes dans la chaîne peptidique de la caséine κ (Fig. 4) a été discutée précédemment en considérant respectivement les deux parties COOH- et NH_2 -terminales correspondant respectivement au caséinomacropéptide [1] et à la paracaséine κ . Sur la Fig. 4, il apparaît clairement que la liaison phénylalanine-méthionine hydrolysée par la présure est située dans une zone hydrophobe de la chaîne, délimitée par deux groupes de résidus basiques.

Récemment, Fiat *et al.* [8] ont déterminé la structure d'un phosphoglycopeptide isolé de la caséine κ . Malheureusement, en nous référant à l'enchaînement des acides aminés proposé, il ne nous a pas été possible de la localiser avec certitude dans la chaîne peptidique de la caséine κ . Rappelons que les motifs glucidiques sont localisés dans la

région 124–155 d'après les résultats de Delfour *et al.* [32] et Pujolle *et al.* [41].

La connaissance du squelette peptidique de la caséine κ devrait faciliter l'étude de la localisation des motifs glucidiques dans les chaînes glycosylées, et permettre de mieux comprendre le rôle que joue cette protéine dans l'édification des micelles et la stabilité de celles-ci.

Le travail qui s'achève par le présent article doit beaucoup à l'excellente collaboration technique de Joël Uro. Les analyses au spectromètre de masse ont été effectuées par M. B. C. Das, dans le Service du Professeur Lederer, à l'Institut de Chimie des Substances Naturelles, C.N.R.S., Gif-sur-Yvette. Nous remercions vivement M. Das de son aide.

ANNEXES

Les figures et tableaux suivants ont été déposés aux Archives originales du Centre de Documentation du C.N.R.S., 26 Rue Boyer, F-75020 Paris, France, où ils peuvent être commandés sous forme de microfiches ou de photocopies, Numéro de référence de dépôt: A.O. — 523.

Annexe, Fig. 1. Diagramme d'élution du surnageant de l'hydrolysats tryptique de la SCM-paracaséine κ .

Annexe, Fig. 2. Diagramme d'élution de l'hydrolysats peptique du peptide tryptique T14.

Annexe, Fig. 3. Diagramme d'élution d'un hydrolysats chymotrypsique de 30 μ moles du peptide κ BCN1 maléylé.

Annexe, Tableau 1. Etude du peptide tryptique T1 de la SCM-paracaséine κ B.

Annexe, Tableau 2. Etude du peptide tryptique T2 de la SCM paracaséine κ B.

Annexe, Tableau 3. Etude du peptide tryptique T3 de la SCM-paracaséine κ B.

Annexe, Tableau 4. Etude des peptides tryptiques T4, T5 et T6 de la SCM-paracaséine κ B.

Annexe, Tableau 5. Etude des peptides tryptiques T7, T8, T9 et T10 de la SCM-paracaséine κ B.

Annexe, Tableau 6. Etude des peptides tryptiques T11, T12 et T13 de la SCM-paracaséine κ B.

Annexe, Tableau 7. (Suite 1 et 2). Etude des fragments peptiques P1 $\rightarrow$ P7 du peptide tryptique T14.

Annexe, Tableau 8. (Suite 1, 2, 3 et 4). Etude des peptides chymotrypsiques C1 $\rightarrow$ C21 du peptide κ BCN1 maléylé.

Annexe, Tableau 9. Etude du peptide "NBS" de la paracaséine κ B.

BIBLIOGRAPHIE

- Mercier, J.-C., Uro, J., Ribadeau Dumas, B. & Grosclaude, F. (1972) *Eur. J. Biochem.* 27, 535.
- Brignon, G., Mercier, J.-C., Ribadeau Dumas, B. & Das, B. C. (1972) *FEBS Lett.* 27, 301.
- Bigelow, C. C. (1967) *J. Theor. Biol.* 16, 187.
- Herskovits, T. T. (1966) *Biochemistry*, 5, 1018.
- Blombäck, B. (1967) *Methods Enzymol.* 11, 398.
- Grosclaude, F., Mahé, M. F., Mercier, J.-C. & Ribadeau Dumas, B. (1972) *Ann. Génét. Sél. Anim.* 4, 515.
- Jollès, J., Alais, C. & Jollès, P. (1970) *Helv. Chim. Acta*, 53, 1918.
- Fiat, A. M., Alais, C. & Jollès, P. (1972) *Eur. J. Biochem.* 27, 408.
- Kim, Y. K., Yaguchi, M. & Rose, D. (1969) *J. Dairy Sci.* 52, 316.
- Schmidt, D. G. (1964) *Biochim. Biophys. Acta*, 90, 411.
- Butler, P. J. G., Harris, J. I., Hartley, B. S. & Leberman, R. (1967) *Biochem. J.* 103, 78.

12. Grosclaude, F., Mercier, J.-C. & Ribadeau Dumas, B. (1970) *Eur. J. Biochem.* 14, 98.
13. Woychik, J. H. (1965) *Arch. Biochem. Biophys.* 109, 542.
14. Sprössler, B., Heilmann, H.-D., Grampp, E. & Uhlig, H. (1971) *Hoppe-Seyler's Z. Physiol. Chem.* 352, 1524.
15. Mercier, J.-C., Grosclaude, F. & Ribadeau Dumas, B. (1970) *Eur. J. Biochem.* 14, 108.
16. Ribadeau Dumas, B., Grosclaude, F. & Mercier, J.-C. (1971) *Eur. J. Biochem.* 18, 252.
17. Ribadeau Dumas, B., Brignon, G., Grosclaude, F. & Mercier, J.-C. (1972) *Eur. J. Biochem.* 25, 505.
18. Spies, J. R. (1967) *Anal. Chem.* 39, 1412.
19. Ramachandran, L. K. & Witkop, B. (1964) *Biochemistry*, 3, 1603.
20. Fraenkel-Conrat, H. & Chung Ming Tsung (1967) *Methods Enzymol.* 11, 151.
21. Mercier, J.-C., Grosclaude, F. & Ribadeau Dumas, B. (1970) *Eur. J. Biochem.* 16, 439.
22. Thomas, D. W. (1969) *FEBS Lett.* 25, 253.
23. Thomas, D. W., Das, B. C., Géro, S. D. & Lederer, E. (1968) *Biochem. Biophys. Res. Commun.* 32, 519.
24. Kalan, E. B. & Woychik, J. H. (1965) *J. Dairy Sci.* 48, 1423.
25. De Koning, P. J., Van Rooijen & Kok, A. (1966) *Biochem. Biophys. Res. Commun.* 24, 616.
26. Schmidt, D. G., Both, P. & De Koning, P. J. (1966) *J. Dairy Sci.* 49, 776.
27. Woychik, J. H., Kalan, E. B. & Noelken, M. E. (1966) *Biochemistry*, 5, 2276.
28. Delfour, A., Jollès, J., Alais, C. & Jollès, P. (1965) *Biochem. Biophys. Res. Commun.* 19, 452.
29. Malpress, F. H. & Seid-Akhavan, M. (1966) *Biochem. J.* 101, 764.
30. Hill, R. J., Naughton, M. A. & Wake, R. G. (1970) *Biochim. Biophys. Acta*, 200, 267.
31. Kuwata, T., Niki, R. & Arima, S. (1969) *J. Agric. Chem. Soc. Jap.* 43, 183.
32. Delfour, A., Alais, C. & Jollès, P. (1966) *Chimia*, 20, 148.
33. Jollès, P., Alais, C. & Jollès, J. (1962) *Arch. Biochem. Biophys.* 98, 56.
34. Jollès, P., Alais, C. & Jollès, J. (1963) *Biochim. Biophys. Acta*, 69, 511.
35. De Koning, P. J. (1967) Thèse, Université d'Amsterdam.
36. Jollès, J., Alais, C. & Jollès, P. (1968) *Biochim. Biophys. Acta*, 168, 591.
37. Hill, R. J. & Wake, R. G. (1969) *Biochim. Biophys. Acta*, 175, 419.
38. Leslie, R. B., Irons, L. & Chapman, D. (1969) *Biochim. Biophys. Acta*, 188, 237.
39. Thomas, M. A. W. (1972). *Symposium sur les Caséines*, Ede, Pays-Bas.
40. Mercier, J.-C., Grosclaude, F. & Ribadeau Dumas, B. (1971) *Eur. J. Biochem.* 23, 41.
41. Pujolle, J., Ribadeau Dumas, B., Garnier, J. & Pion, R. (1966) *Biochem. Biophys. Res. Commun.* 25, 285.

J.-C. Mercier, G. Brignon et B. Ribadeau-Dumas
Laboratoire de Recherches sur les Protéines, C.N.R.Z.
Institut National de la Recherche Agronomique
Domaine de Vilvert, F-78350 Jouy-en-Josas, France